OM nucleic - nucleic search, using sw model

Run on: November 20, 1999, 20:38:59 ; Search time 48.9 Seconds
(without alignments)
9690.466 Million cell updates/sec

Title: US-09-126-945-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUCDX

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	78.4	4.1	2667	1	T47198	Human ETS2 repress
2	78.2	4.1	1116	1	V38354	Transcription fact
3	78.2	4.1	1907	1	V38355	Transcription fact
4	78	4.1	2073	1	Q55149	Adenovirus E1A-F g
5	78	4.1	2064	1	T37087	E1AF matrix metall
6	76.6	4.0	1920	1	V40701	Human ELF3 protein
7	76.4	4.0	1800	1	Q91769	Coding sequence of
8	74.2	3.9	2938	1	Q50644	Human Hum-Fli-1 ge
9	74.2	3.9	1890	1	Q50662	Human Hum-Fli-1 ge
10	73.4	3.9	2410	1	V32688	Polyomavirus PEA3
11	68.6	3.6	7752	1	V38365	Mouse ESX transcri
c 12	60.6	3.2	852	1	V58521	Prostate tumour sp
c 13	60.6	3.2	852	1	V61167	cDNA sequence of p
14	60.4	3.2	2975	1	V15529	Human lymphoid-spe
15	60.4	3.2	3240	1	V15530	Human lymphoid-spe
16	60	3.2	2544	1	T47200	Mouse ETS2 repress
17	59.6	3.1	1604	1	V20465	Human c-ets oncoge
18	53.4	2.8	1447	1	X26551	DNA encoding chick
19	53.4	2.8	1528	1	X26552	DNA encoding chick
20	49	2.6	1364	1	V20472	Human c-spi oncoge
21	46.2	2.4	1913	1	Q15015	Nuclear factor C/E
22	45.8	2.4	15079	1	Q91580	S. clavuligerus cl
23	45.4	2.4	2386	1	Q63477	Micrococcal putres
24	45.2	2.4	15872	1	T68715	Streptomyces venez
25	45.2	2.4	43280	1	T80413	Tylactone synthase
26	44.6	2.4	2544	1	T67066	Human alpha3(IX) c
c 27	43	2.3	985	1	V44439	Mycobacterium tube
c 28	43	2.3	985	1	V64548	M. tuberculosis im
29	42.6	2.2	2338	1	Q14183	N.clavipes draglin
30	42.6	2.2	12001	1	Q76213	HSV L/ST region. H
31	42.6	2.2	2338	1	V23249	Nephila clavipes s
c 32	42.6	2.2	4626	1	X33449	Oryza sativa L. pi
33	42	2.2	2004	1	T85356	Nephila clavipes s
c 34	41.6	2.2	2568	1	T88419	Human glial cell l
c 35	41.6	2.2	1223	1	V00246	Human Ret ligand r
c 36	41.6	2.2	1682	1	V00247	Human Ret ligand r
c 37	41.6	2.2	2568	1	V99310	Glial cell-line de
c 38	41.6	2.2	1927	1	V99315	GDNFR-alpha cDNA c
c 39	41.6	2.2	1929	1	V99316	GDNFR-alpha cDNA c
40	41.4	2.2	53789	1	V21187	Amycolatopsis medi
c 41	41.2	2.2	738	1	X39714	Gastric cancer ass
42	41.2	2.2	927	1	X39716	Gastric cancer ass
43	41	2.2	1985	1	Q37480	Sequence encoding
44	40.6	2.1	23666	1	Q10190	Cephalosporin anti
45	40.6	2.1	44377	1	T78508	Platenolide syntha

ALIGNMENTS

RESULT 1
T47198
ID T47198 standard; cDNA; 2667 BP.
AC T47198;
DT 06-APR-1997 (first entry)
DE Human ETS2 repressor factor (ERF) cDNA.
KW ETS2 repressor factor; ERF; transcriptional repressor;
KW tumour suppressor; tumour; cancer; oncogene; gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 123. .1769
FT /*tag= a
FT polya_signal 2644. .2649
FT /*tag= b
PN WO9639517-A1.
PD 12-DEC-1996.
PF 04-JUN-1996; U10177.
PR 05-JUN-1995; US-469412.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Athanasiou MA, Beal GJ, Blair DG, Fisher RJ, Mavrothalassitis GJ;
PI Sgouras D N;
DR WPI; 97-043139/04.
DR P-PSDB; W07700.
PT New DNA encoding ETS2 repressor factor - useful for reducing
PT tumourigenicity, esp. oncogene associated tumour cells
PS Claim 3; Page 63-65; 101pp; English.
CC A cDNA clone (T47198) codes for human ETS2 repressor factor (ERF)
CC (W07700). It was isolated from a K562 cDNA library using the H1
CC site of the ETS2 promoter as probe. A related clone (T47199)
CC coding for an alternatively spliced ERF (W07701) was also isolated.
CC The ERF gene, which maps to chromosome 19, q1.2-1.3, is the first
CC member of the ets family to be identified as a transcriptional
CC repressor in mammalian cells. It can be used to suppress or repress
CC transcription and to elucidate transcription process and regulation.
CC The ERF gene also has tumour suppresor activity and can be used to
CC reduce ets-dependent tumourigenicity associated with v-mos, c-met,
CC tpr-met, Ha-ras and gag-myb-ets oncogenes. The cDNA can be cloned
CC for expression of the ERF polypeptide in host cells.
SQ Sequence 2667 BP; 456 A; 876 C; 825 G; 510 T;

Query Match 4.1%; Score 78.4; DB 1; Length 2667;
Best Local Similarity 59.8%; Pred. No. 1.3e-07;
Matches 152; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

```
Qy   1124 gacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttcctcaag 1183
          | |  | ||   || || ||      |    ||  ||||| |||||||| ||  ||  |
Db    165 GCCTACAAGCCAGAGTCGTCCCCTGGCTCAAGGCAGATCCAGCTGTGGCACTTTATCCTG 224

Qy   1184 gagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaaggagaag 1243
          ||| |||| |  |||    |   | | |||  | |      ||  |    || | | | |
Db    225 GAGCTGCTGCGGAAGGAGGAGTACCAGGGCGTCATTGCCTGGCAGGGGGACTACGGGGAA 284

Qy   1244 ggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaac 1303
            | || |||||     ||| | |   ||||||||||||||||||||| | ||||||  |
Db    285 TTCGTCATCAAA------GACCCTGATGAGGTGGCCCGGCTGTGGGGCGTTCGCAAGTGC 338

Qy   1304 cgtcccgccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggc 1363
           |||  ||||| ||||||||||||||||||  || | ||| | || || || ||| ||
Db    339 AAGCCCCAGATGAATTACGACAAGCTGAGCCGGGCCCTGCGCTATTACTATAACAAGCGC 398

Qy   1364 atcatccggaagcc 1377
          ||  | |  ||| |
Db    399 ATTCTGCACAAGAC 412
```

```
RESULT   2
V38354
ID   V38354 standard; cDNA; 1116 BP.
AC   V38354;
DT   09-NOV-1998  (first entry)
DE   Transcription factor ESX gene open reading frame.
KW   ESX; epithelial-restricted with serine box; transcription factor;
KW   ETS; human; epithelial cancer; breast cancer; adenocarcinoma;
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   exon            1. .163
FT                   /*tag= a
FT                   /number= 2
FT   exon            164. .385
FT                   /*tag= b
FT                   /number= 3
FT   exon            386. .478
FT                   /*tag= c
FT                   /number= 3
FT   exon            479. .688
FT                   /*tag= d
FT                   /number= 5
FT   exon            689. .805
FT                   /*tag= e
FT                   /number= 6
FT   exon            806. .1001
FT                   /*tag= f
FT                   /number= 8
FT   exon            1002. .1111
FT                   /*tag= g
FT                   /number= 9
FT   misc_feature    1. .189
FT                   /*tag= h
FT                   /product= first variable region
FT                   /note= "Claim 28"
FT   misc_feature    190. .309
FT                   /*tag= i
FT                   /product= pointed region
FT   misc_feature    310. .561
FT                   /*tag= j
FT                   /product= second variable region
FT                   /note= "Claim 8"
FT   misc_feature    562. .714
FT                   /*tag= k
FT                   /product= serine-rich region
FT   misc_feature    715. .819
FT                   /*tag= l
FT                   /product= third variable region
FT                   /note= "Claim 28"
FT   misc_feature    820. .1062
FT                   /*tag= m
FT                   /product= Ets DNA binding domain
FT   misc_feature    1063. .1113
FT                   /*tag= n
FT                   /product= fourth variable region
FT                   /note= "Claim 28"
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60677-79.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
PT   useful for diagnosis and treatment of cancer
PS   Claim 3; Fig 1; 120pp; English.
CC   This is the open reading frame of the newly isolated ETX gene that
CC   codes for a novel transcription factor (see W60677) associated
CC   with the aetiology of cancers, including epithelial cancers.
CC   The gene is a member of the ETS family and is designated ETX for
CC   epithelial-restricted with serine box.  The gene is located at
CC   chromosome 1q32, a region amplified in 50% of early breast cancers.
CC   ESX is an early indicator of breast cancer, being expressed at the
CC   ductal carcinoma in situ stage.  Dysfunction of the ESX gene that
CC   results in increased expression is indicative of epithelial cancer
CC   (especially breast cancer but also gastric, ovarian and lung
CC   adenocarcinoma) and of unfavourable prognosis.  The ESX open
CC   reading frame was deduced from a cDNA clone of foetal liver-spleen
CC   origin that was identified in an expressed sequence tag database
CC   search using human Ets-2, and from a clone obtained by 5'RACE of
CC   placental polyA mRNA.  Antibodies, antisense nucleic acids or
CC   inactive ESX mutant proteins that inhibit activity of ESX can be
CC   used to inhibit growth and proliferation of neoplastic cells,
CC   particularly cancer.  ESX polypeptides, polynucleotides and
CC   antibodies are also used to detect ESX, particularly for diagnosis
CC   and monitoring of cancer.  Polynucleotides may also be used to
CC   detect gross abnormalities in chromosome 1, e.g. for in utero
CC   screening for predisposition to cancer, and to generate ESX protein
CC   in vivo from gene therapy vectors.  ESX protein, genes and cDNA
CC   can also be used to screen for specific modulators, especially
CC   potential anticancer agents.
SQ   Sequence  1116 BP;  255 A;  322 C;  335 G;  204 T;

  Query Match             4.1%;  Score 78.2;  DB 1;  Length 1116;
  Best Local Similarity   59.8%;  Pred. No. 1.2e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
           || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    816 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 875

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
           ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |.    |||||
Db    876 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 935

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
           ||  || ||||||      || ||      |  ||||| |||||| ||||||||||  |
Db    936 CCAACTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 995

Qy   1339 catccgccagtattacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
           |||  |   ||| ||| | ||  |    |||| | | |  |       || ||| |||||
Db    996 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1052

Qy   1399 ctaccagttcg 1409
           |||| |||| |
Db   1053 CTACAAGTTTG 1063


RESULT   3
V38355
ID   V38355 standard; cDNA; 1907 BP.
AC   V38355;
DT   09-NOV-1998  (first entry)
DE   Transcription factor ESX cDNA (ORF and untranslated regions).
KW   ESX; epithelial-restricted with serine box; transcription factor;
KW   ETS; human; epithelial cancer; breast cancer; adenocarcinoma;
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   5'UTR           1. .95
FT                   /*tag= a
FT   CDS             96. .1206
FT                   /*tag= b
FT                   /note= "Claim 3"
FT   3'UTR           1207. .1907
FT                   /*tag= c
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
```

```
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60677.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
PT   useful for diagnosis and treatment of cancer
PS   Disclosure; Page 83; 120pp; English.
CC   This cDNA sequence comprises the open reading frame (see also
CC   V38354) and 5' and 3' untranslated regions of the newly isolated
CC   ETX gene that codes for a novel human transcription factor (see
CC   W60677) associated with the aetiology of cancers. The ETX gene
CC   is a member of the ETS family and is designated ETX for epithelial
CC   restricted with serine box. The gene is located at chromosome
CC   1q32, a region amplified in 50% of early breast cancers. ESX is an
CC   early indicator of breast cancer, being expressed at the ductal
CC   carcinoma in situ stage. Dysfunction of the ESX gene that results
CC   in increased expression is indicative of epithelial cancer
CC   (especially breast cancer but also gastric, ovarian and lung
CC   adenocarcinoma) and of unfavourable prognosis. The ESX open
CC   reading frame was deduced from a cDNA clone of foetal liver-spleen
CC   origin that was identified in an expressed sequence tag database
CC   search using human Ets-2, and from a clone obtained by 5'RACE of
CC   placental polyA mRNA. Antibodies, antisense nucleic acids or
CC   inactive ESX mutant proteins that inhibit activity of ESX can be
CC   used to inhibit growth and proliferation of neoplastic cells,
CC   particularly cancer. ESX polynucleotides, polypeptides and
CC   antibodies are also used to detect ESX, particularly for diagnosis
CC   and monitoring of cancer. ESX polynucleotides may also be used to
CC   detect gross abnormalities in chromosome 1, e.g. for in utero
CC   screening for predisposition to cancer, and to generate ESX protein
CC   in vivo from gene therapy vectors. ESX protein, genes and cDNA can
CC   also be used to screen for specific modulators, especially
CC   potential anticancer agents.
SQ   Sequence   1907 BP;    427 A;    557 C;    533 G;    390 T;

  Query Match             4.1%;  Score 78.2;  DB 1;  Length 1907;
  Best Local Similarity   59.8%;  Pred. No. 1.3e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcctt 1218
           || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    911 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 970

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
           ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    971 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 1030

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
           ||  || ||||||      || ||      |  ||||| |||||| ||||||||||  |
Db   1031 CCAACTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 1090

Qy   1339 catccgccagtattacaagaaggggcatcatccggaagccagacatctcccagcgcctcgt 1398
           |||  |  ||| ||| | ||  |  |||| | |  |        || ||| |||||
Db   1091 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1147

Qy   1399 ctaccagttcg 1409
           |||| |||| |
Db   1148 CTACAAGTTTG 1158


RESULT   4
Q55149
ID   Q55149 standard; cDNA to mRNA; 2073 BP.
AC   Q55149;
DT   11-JUL-1994  (first entry)
DE   Adenovirus E1A-F gene.
KW   Adenovirus; cancer; ets oncogene; HeLa cell; enhancer core sequence;
KW   methylation; ds.
OS   Human adenovirus.
FH   Key             Location/Qualifiers
FT   cds             844. .1311
FT                   /*tag= a
FT                   /note= "Claimed sequence"
PN   J05328975-A.
PD   14-DEC-1993.
PF   02-JUN-1992; 165453.
PR   02-JUN-1992; JP-165453.
PA   (TAKI ) TAKARA SHUZO CO LTD.
DR   WPI; 94-021923/03.
DR   P-PSDB; R45451.
PT   Novel E1A-F gene - for production of adenovirus E1A-F and cancer
PT   research
PS   Claim 1; Page 6; 7pp; Japanese.
CC   The adenovirus E1A-F gene contains a 473bp open reading frame. The
CC   clone comprising the coding sequence was isolated by screening
CC   a HeLa cell cDNA library.
SQ   Sequence   2073 BP;    458 A;    635 C;    561 G;    418 T;

  Query Match             4.1%;  Score 78;  DB 1;  Length 2073;
  Best Local Similarity   76.2%;  Pred. No. 1.5e-07;
  Matches   96;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
           |||||  |    ||  | |   |||| ||| |||| ||||||||||  |||||||| ||
Db    961 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1020

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
           ||||||||  ||||||||||||||||||||||  ||||  | |||||  |||| ||||||||
Db   1021 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1080

Qy   1370 cggaag 1375
           | ||||
Db   1081 CAGAAG 1086


RESULT   5
T37087
ID   T37087 standard; cDNA to mRNA; 2064 BP.
AC   T37087;
DT   25-APR-1997  (first entry)
DE   E1AF matrix metalloproteinase regulator, cDNA.
KW   E1AF; matrix; metalloproteinase; regulator; infiltration; cancer;
KW   metastasis; cell; control; antisense; decoy; DNA binding region;
KW   target DNA; ribosome; induction; diagnosis; detection; treatment;
KW   mammary cancer; fibrosarcoma; osteosarcoma; lung cancer; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .1389
FT                   /*tag= a
PN   WO9624379-A1.
PD   15-AUG-1996.
PF   09-JAN-1996; J00016.
PR   08-FEB-1995; JP-020173.
PA   (TAKI ) TAKARA SHUZO CO LTD.
PI   Fujinaga K,  Higashino F,  Yoshida K;
DR   WPI; 96-384227/38.
DR   P-PSDB; W00167.
PT   Control of cancer cell infiltration by E1AF gene expression
PT   regulation  - also diagnosis of cancer by detection of E1AF gene
PT   expression products
PS   Example 5; Pages 38-42; 65pp; Japanese.
CC   The present sequence encodes the E1AF protein, which is a matrix
CC   metalloproteinase regulator. The infiltration and metastasis of
CC   cancer cells can be controlled by regulating the expression and
CC   expression products of the E1AF gene. This may be accomplished by
CC   inducing antisense DNA or RNA for the E1AF gene, a decoy gene
CC   expressing the DNA binding region of the E1AF protein, the target
CC   DNA for the DNA binding region of the E1AF protein or ribosomes
CC   corresponding to the E1AF gene mRNA. Cancer can be diagnosed by
CC   detecting E1AF gene expression products, e.g. E1AF protein or mRNA.
CC   These methods may be used in the treatment and diagnosis of cancer,
```

```
CC   e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc. .
SQ   Sequence   2064 BP;    427 A;    648 C;    569 G;    420 T;

  Query Match             4.1%;  Score 78;  DB 1;  Length 2064;
  Best Local Similarity   76.2%;  Pred. No. 1.5e-07;
  Matches   96;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
           |||||  |    ||  | |   |||| ||| |||| ||||||||||  |||||||| ||
Db   1039 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1098

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
           |||||||| ||||||||||||||||||||||  ||||  | |||||  |||| ||||||||
Db   1099 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1158

Qy   1370 cggaag 1375
           | ||||
Db   1159 CAGAAG 1164


RESULT   6
V40701
ID   V40701 standard; DNA; 1920 BP.
AC   V40701;
DT   21-SEP-1998  (first entry)
DE   Human ELF3 protein coding sequence.
KW   ELF3; human; cancer; epithelial cancer; breast cancer; prostate cancer;
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             115. .1230
FT                   /*tag= a
FT                   /product= ELF3
PN   EP-839908-A2.
PD   06-MAY-1998.
PF   23-OCT-1997; 308450.
PR   15-NOV-1996; US-746789.
PR   31-OCT-1996; US-028791.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (MONU ) UNIV MONASH.
PI   Debouck C, Kola I, Tymms MJ;
DR   WPI; 98-242674/22.
DR   P-PSDB; W57834.
PT   New isolated polynucleotide and e.g. DNA, agonists and antagonists -
PT   used to diagnose and treat cancer, e.g. prostate, breast, lung or
PT   epithelial tumours
PS   Claim 4; Fig 1A; 34pp; English.
CC   This sequence encodes the human ELF3 protein of the invention. ELF3 is
CC   used to treat patients requiring ELF3 by providing DNA encoding the
CC   polypeptide and expressing the polypeptide in vivo. ELF3 is used to
CC   diagnose a disease or a susceptibility to a disease related to its
CC   expression by determining a mutation in the nucleic acid sequence
CC   encoding it. The presence ELF3 in a sample derived from a host can be
CC   detected. Antagonists of ELF3 are used to treat patients requiring
CC   inhibition of ELF3 polypeptide. ELF3 can be used in the early diagnosis
CC   and treatment of cancer, particularly epithelial cancer such as breast
CC   and prostate cancer.
SQ   Sequence   1920 BP;    422 A;    561 C;    550 G;    387 T;

  Query Match             4.0%;  Score 76.6;  DB 1;  Length 1920;
  Best Local Similarity   59.4%;  Pred. No. 2.8e-07;
  Matches  149;  Conservative    0;  Mismatches   99;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
           || ||||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    930 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 989

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
           ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    990 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 1049

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
           ||  || |||||        || ||      |  ||||| |||||| ||||||||||  |
Db   1050 CCAACTATGGGGGCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 1109

Qy   1339 catccgccagtattacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
           |||  |   ||| ||| | ||  |    |||| | | |         || ||| |||||
Db   1110 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1166

Qy   1399 ctaccagttcg 1409
           |||| |||| |
Db   1167 CTACAAGTTTG 1177


RESULT   7
Q91769
ID   Q91769 standard; cDNA to mRNA; 1800 BP.
AC   Q91769;
DT   09-FEB-1996  (first entry)
DE   Coding sequence of PEA3-beta, an ETS transcription factor.
KW   Transcription factor; probe; reverse transcription; PCR; primer;
KW   expression vector; E.coli; COS cell; ras; cancer cell multiplication;
KW   polyoma virus; transformation; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .1563
FT                   /*tag= a
FT                   /product= PEA3-beta transcription factor
FT   cds             250. .252
FT                   /*tag= b
FT                   /transl_except= seq: CTC  a.a.:Pro
FT   cds             376. .378
FT                   /*tag= c
FT                   /transl_except= seq: AAC  a.a.: Lys
FT   cds             1498. .1500
FT                   /*tag= d
FT                   /transl_except= seq: ACC  a.a.: Ser
PN   J07145197-A.
PD   06-JUN-1995.
PF   25-NOV-1993; 295393.
PR   25-NOV-1993; JP-295393.
PA   (EISA ) EISAI CO LTD.
PA   (HIRA/) HIRANO T.
DR   WPI; 95-237197/31.
DR   P-PSDB; R78185.
PT   ETS transcription factor activated by ras - may be used in the study
PT   of cancer cell proliferation and the proliferation of the polyoma
PT   virus
PS   Claim 2; Page 7-9; 9pp; Japanese.
CC   The nucleotide sequence of the novel ETS transcription factor family
CC   member - PEA3-beta.  The gene was isolated from a HepG2 cell line cDNA
CC   library.  The probe for the screening was prepared by reverse
CC   transcription on HepG2 mRNA followed by PCR using primers Q91770-1,
CC   to produce a probe of 170-200 bp.  The screening isolated the full
CC   length sequence of the transcription factor.  The gene was inserted into
CC   the expression vectors pBluescript KS and pcDV1 for expression of the
CC   protein in E.coli and COS7 cells, respectively.  The ETS transcription
CC   factor has specificity for and is activated by ras.  It is useful as a
CC   reagent in studies for the elucidation of the mechanism of cancer cell
CC   multiplication or polyoma virus transformation of cells.
SQ   Sequence   1800 BP;    437 A;    522 C;    435 G;    405 T;

  Query Match             4.0%;  Score 76.4;  DB 1;  Length 1800;
  Best Local Similarity   75.4%;  Pred. No. 3e-07;
  Matches   95;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
           |||||  |    ||  | |   |||| || ||||  |||||||||  |||||||| ||
Db   1186 TTCAAGCTGATAGAACCGGAAGAGGTTGCTCGGCGCTGGGGCATCCAGAAGAACCGGCCA 1245

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
           |||||||||| ||||||||||||||||  |||||  | || ||  | ||||||||||
Db   1246 GCCATGAACTATGACAAGCTGAGCCGCTCTCTCCGCTATTACTATGAAAAGGGCATCATG 1305
```

```
Qy   1370 cggaag 1375
            | ||||
Db   1306 CAGAAG 1311


RESULT   8
Q50644
ID   Q50644 standard; cDNA; 2938 BP.
AC   Q50644;
DT   26-MAY-1994  (first entry)
DE   Human Hum-Fli-1 gene clone BM025.
KW   chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
KW   Ews gene; malignant melanoma; hum-fli-1;
KW   primitive peripheral neuroectodermal tumour; human chromosome 11;
KW   human chromosome 22; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             143. .1501
FT                   /*tag= a
FT                   /product= HUM-FLI-1
FT   polya_signal    2908. .2913
FT                   /*tag= b
PN   WO9323549-A.
PD   25-NOV-1993.
PF   19-MAY-1993; F00494.
PR   20-MAY-1992; FR-006123.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PI   Aurias A,  Delattre O,  Desmaze C,  Melot T,  Peter M;
PI   Plougastel B, Thomas G,  Zucman J;
DR   WPI; 93-386580/48.
DR   P-PSDB; R44556.
PT   New nucleic acid of EWS gene and its hybrid(s) - contg. gene
PT   sequence involved in chromosomal trans-location, also derived
PT   mRNA, probes, fusion proteins etc., for diagnosis and treatment
PT   of Ewing sarcoma and melanoma
PS   Disclosure; Fig 7; 123pp; French.
CC   The probe 11RR1 was used to screen a human marrow cDNA library
CC   (Clontech cat.# HL1058). The clone BM025 was identified and
CC   sequenced. It represents the entire coding region together with
CC   5'- and 3'-UTRs of the Hum-Fli-1 gene.
SQ   Sequence   2938 BP;    847 A;    692 C;    675 G;    724 T;


  Query Match             3.9%;  Score 74.2;  DB 1;  Length 2938;
  Best Local Similarity   62.9%;  Pred. No. 9e-07;
  Matches  134;  Conservative    0;  Mismatches   73;  Indels    6;  Gaps    1;

Qy   1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
             ||  ||||| |||||||| ||||||  |||| ||||    | |    |  || |     || |
Db    979 GCAGATCCAGCTGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTG 1038

Qy   1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggt 1275
              |||    ||  |    | ||| | | ||       ||||||||    |||| | |    ||||
Db   1039 TATCACCTGGGAGGGGACCAACGGGGAG------TTCAAAATGACGGACCCCGATGAGGT 1092

Qy   1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccg 1335
             |||| |||  ||||||    || || | |   |||  |||||| ||||||||||||||||
Db   1093 GGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAAGCTGAGCCG 1152

Qy   1336 ctccatccgccagtattacaagaagggcatcat 1368
              || ||||  | || ||  | ||    ||| ||
Db   1153 GGCCCTCCGTTATTACTATGATAAAAACATTAT 1185


RESULT   9
Q50662
ID   Q50662 standard; DNA; 1890 BP.
AC   Q50662;
DT   26-MAY-1994  (first entry)
DE   Human Hum-Fli-1 gene (genomic DNA).
KW   chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
KW   Ews gene; malignant melanoma; hum-fli-1;
KW   primitive peripheral neuroectodermal tumour; human chromosome 11;
KW   human chromosome 22; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   exon            1. .39
FT                   /*tag= a
FT                   /number= 1
FT                   /codon_start= 22. .24
FT   intron          40. .99
FT                   /*tag= b
FT                   /number= 1
FT   exon            100. .311
FT                   /*tag= c
FT                   /number= 2
FT   intron          312. .371
FT                   /*tag= d
FT                   /number= 2
FT   exon            372. .526
FT                   /*tag= e
FT                   /number= 3
FT   intron          527. .586
FT                   /*tag= f
FT                   /number= 3
FT   exon            587. .790
FT                   /*tag= g
FT                   /number= 4
FT   intron          791. .850
FT                   /*tag= h
FT                   /number= 4
FT   exon            851. .916
FT                   /*tag= i
FT                   /number= 5
FT   intron          917. .976
FT                   /*tag= j
FT                   /number= 5
FT   exon            977. .1042
FT                   /*tag= k
FT                   /number= 6
FT   intron          1043. .1102
FT                   /*tag= l
FT                   /number= 6
FT   exon            1103. .1162
FT                   /*tag= m
FT                   /number= 7
FT   intron          1163. .1222
FT                   /*tag= n
FT                   /number= 7
FT   exon            1223. .1270
FT                   /*tag= o
FT                   /number= 8
FT   intron          1271. .1330
FT                   /*tag= p
FT                   /number= 8
FT   exon            1331. .1890
FT                   /*tag= q
FT                   /number= 9
PN   WO9323549-A.
PD   25-NOV-1993.
PF   19-MAY-1993; F00494.
PR   20-MAY-1992; FR-006123.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PI   Aurias A,  Delattre O,  Desmaze C,  Melot T,  Peter M;
PI   Plougastel B, Thomas G,  Zucman J;
DR   WPI; 93-386580/48.
DR   P-PSDB; R44556.
PT   New nucleic acid of EWS gene and its hybrid(s) - contg. gene
PT   sequence involved in chromosomal trans-location, also derived
PT   mRNA, probes, fusion proteins etc., for diagnosis and treatment
PT   of Ewing sarcoma and melanoma
PS   Disclosure; Fig 7 and Fig 13; 123pp; French.
CC   The intron-exon junctions of the human Ews gene and the Hum-Fli-1
CC   gene have been sequenced (see Q50646 and Q50662, respectively).
CC   The different fusion products which could be formed by fusing
```

```
CC   exons from the two genes, as happens after specific chromosomal
CC   translocations, can be predicted (see Q50671-Q50678). The sequences
CC   at fusion junctions of other observed translocations are given in
CC   Q50679-Q50683). NOTE: the 1890 residue sequence Q50662 does not
CC   appear in the specification; it is a combination of the Hum-Fli-1
CC   cDNA sequence (Fig 7, Q50644) and the individual intron sequences
CC   with their intron-exon junctions (Fig 13).
SQ   Sequence   1890 BP;    457 A;    561 C;    458 G;    414 T;

  Query Match             3.9%;  Score 74.2;  DB 1;  Length 1890;
  Best Local Similarity   62.9%;  Pred. No. 8.5e-07;
  Matches  134;  Conservative    0;  Mismatches   73;  Indels    6;  Gaps    1;

Qy    1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
           ||  ||||| |||||||| |||||| |||| ||||   | |   |  || |    || |
Db    1338 GCAGATCCAGCTGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTG 1397

Qy    1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagccccaggt 1275
             |||   ||  |   | ||| | | ||       |||||||   |||| | |   ||||
Db    1398 TATCACCTGGGAGGGGACCAACGGGGAG------TTCAAAATGACGGACCCCGATGAGGT 1451

Qy    1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccg 1335
           |||| |||  ||||||   || || | |   |||  |||||| ||||||||||||||||
Db    1452 GGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAAGCTGAGCCG 1511

Qy    1336 ctccatccgccagtattacaagaagggcatcat 1368
             || ||||  | || ||  | ||   ||| ||
Db    1512 GGCCCTCCGTTATTACTATGATAAAAACATTAT 1544


RESULT  10
V32688
ID   V32688 standard; cDNA; 2410 BP.
AC   V32688;
DT   20-OCT-1998  (first entry)
DE   Polyomavirus PEA3 cDNA.
KW   Polyomavirus enhancer activator; PEA3; tumour; suppressor; inhibitor;
KW   transformation; HER-2; neu promoter; metastasis; cancer; ss.
OS   Polyomavirus.
FH   Key             Location/Qualifiers
FT   CDS             126. .1793
FT                   /*tag=  a
FT                   /product= PEA3
FT                   /note= "polyomavirus enhancer activator"
PN   WO9830585-A2.
PD   16-JUL-1998.
PF   12-JAN-1998; U00880.
PR   10-JAN-1997; US-780835.
PA   (TEXA ) UNIV TEXAS SYSTEM.
PI   Hung M, Xing X;
DR   WPI; 98-399061/34.
DR   P-PSDB; W49010.
PT   Method for repressing transformation of cells - by contacting cell
PT   with polyoma-virus enhancer activator 3, useful for, e.g. treating
PT   or preventing cancer, tumourigenesis and metastasis
PS   Disclosure; Page 69-70; 83pp; English.
CC   This sequence encodes a polyomavirus enhancer activator, PEA3. This
CC   sequence is used in a method for repressing transformation in a cell
CC   which involves contacting the cell with PEA3 to inhibit a transformed
CC   phenotype. This sequence can also be used in a method to suppress the
CC   growth of a tumour in a mammal comprising introducing to the mammal a
CC   PEA3-encoding nucleic acid where the expression of PEA3 in the mammal
CC   results in a decrease in the growth rate of the tumour. PEA3 can regulate
CC   the HER-2/neu promoter by transcriptional repression and acts as a tumour
CC   repressor. The PEA3 can be used for reducing a transforming,
CC   tumourogenic or metastatic potential of a cell. It can be used for the
CC   prevention and treatment of such transformation-driven events as cancer,
CC   tumourigenesis and metastasis.
SQ   Sequence   2410 BP;    521 A;    744 C;    645 G;    500 T;

  Query Match             3.9%;  Score 73.4;  DB 1;  Length 2410;
  Best Local Similarity   59.9%;  Pred. No. 1.3e-06;
  Matches  160;  Conservative    0;  Mismatches  101;  Indels    6;  Gaps    2;

Qy    1109 gaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctg 1168
           || | ||   |  |  | | |||| |   | |     | ||    || | || |||
Db    1308 GAAGGGATTGGAGCTTTCCGGGAGGGGCCACCCTACCAGCGCCGGGGTGCCTTACAACTG 1367

Qy    1169 tggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtgg 1228
           |||||||| ||   || |  || ||  |   ||  |   |||  |  |||||   |||
Db    1368 TGGCAGTTTCT---GGTGGCCCTGCTGGATGACCCAACAAATGCTCATTTCATTGCTTGG 1424

Qy    1229 ctcaacaaggagaagggcatcttcaaaattgaggactcagccccaggtggcccggctgtgg 1288
             |  ||  | |||       || ||| |   ||  | |   |||| ||| |||| |||
Db    1425 ACAGGCCGGGGAATGG---AGTTTAAACTAATTGAACCTGAAGAGGTTGCCAGGCTCTGG 1481

Qy    1289 ggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatccgccag 1348
           || ||||  |||||||| || |||||||| || ||||||||||||||||  | ||  |
Db    1482 GGTATCCAGAAGAACCGGCCAGCCATGAATTATGACAAGCTGAGCCGCTCGCTGCGATAC 1541

Qy    1349 tattacaagaagggcatcatccggaag 1375
           |||||  |||| |||||||| | ||||
Db    1542 TATTATGAGAAAGGCATCATGCAGAAG 1568


RESULT  11
V38365
ID   V38365 standard; cDNA; 7752 BP.
AC   V38365;
DT   09-NOV-1998  (first entry)
DE   Mouse ESX transcription factor gene.
KW   ESX; epithelial-restricted with serine box; transcription factor;
KW   ETS; mouse; epithelial cancer; breast cancer; adenocarcinoma;
KW   diagnosis; therapy; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   promoter        2553. .2889
FT                   /*tag=  a
FT   CDS             3604. .7115
FT   ,               /*tag=  b
FT                   /note= "contains introns"
FT   exon            2900. .3035
FT                   /*tag=  c
FT                   /number= 1
FT   intron          3036. .3606
FT                   /*tag=  d
FT                   /number= 1
FT   exon            3604. .3764
FT                   /*tag=  e
FT                   /number= 2
FT   intron          3765. .4148
FT                   /*tag=  f
FT                   /number= 2
FT   exon            4149. .4373
FT                   /*tag=  g
FT                   /number= 3
FT   intron          4374. .4503
FT                   /*tag=  h
FT                   /number= 3
FT   exon            4504. .4594
FT                   /*tag=  i
FT                   /number= 4
FT   intron          4595. .4782
FT                   /*tag=  j
FT                   /number= 4
FT   exon            4783. .4905
FT                   /*tag=  k
FT                   /number= 5
FT   intron          4906. .5052
FT                   /*tag=  l
FT                   /number= 5
FT   exon            5053. .5142
FT                   /*tag=  m
FT                   /number= 6
```

```
FT   intron          5143. .5284
FT                   /*tag= n
FT                   /number= 6
FT   exon            5285. .5402
FT                   /*tag= o
FT                   /number= 7
FT   intron          5403. .6255
FT                   /*tag= p
FT                   /number= 7
FT   exon            6256. .6450
FT                   /*tag= q
FT                   /number= 8
FT   intron          6451. .6998
FT                   /*tag= r
FT                   /number= 8
FT   exon            6999. .7115
FT                   /*tag= s
FT                   /number= 9
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60680.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
PT   useful for diagnosis and treatment of cancer
PS   Claim 18; Page 85-90; 120pp; English.
CC   This is the murine mETX gene that contains about 2.9 kb of
CC   promoter upstream of about 4.9 kb of DNA incorporating at least
CC   9 exons that specify a transcript of about 2 kb, with exons 2-9
CC   encoding the 371-amino acid mESX protein (see W60680).  The gene
CC   is a member of the ETS family and is designated ETX for
CC   epithelial-restricted with serine box.  In humans, a dysfunction of
CC   the ETX gene (see V38354) that results in increased expression is
CC   indicative of epithelial cancer (especially breast cancer but
CC   also gastric, ovarian and lung adenocarcinoma) and of unfavourable
CC   prognosis.  Antibodies, antisense nucleic acids or inactive ESX
CC   mutant proteins that inhibit activity of human ESX can be used to
CC   inhibit growth and proliferation of neoplastic cells, particularly
CC   cancer.  ESX polypeptides, polynucleotides and antibodies are also
CC   used to detect ESX, particularly for diagnosis and monitoring of
CC   cancer.  They can also be used to screen for specific modulators,
CC   especially potential anticancer agents.  Transgenic animals are
CC   also provided.
SQ   Sequence  7752 BP;    1768 A;    2050 C;    2019 G;    1914 T;

  Query Match             3.6%;  Score 68.6;  DB 1;  Length 7752;
  Best Local Similarity   61.5%;  Pred. No. 1.4e-05;
  Matches  110;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy   1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
           ||||||||| ||||  ||   ||  | ||| || | ||| |   | | |   || ||||
Db   6271 CACCTGTGGGAGTTTATCCGAGACATCCTAATCCACCCCGAGCTCAACGAAGGCCTCATG 6330

Qy   1223 aggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccgg 1282
           | ||||   |||  | |  ||||  | |||||  ||     |||||     |||||||
Db   6331 AAGTGGGAGAACCGGCACGAGGGTGTGTTCAAGTTTCTTCGCTCAGAGGCCGTGGCCCAA 6390

Qy   1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccat 1341
           || |||||||      |||||       |  ||||| ||| || ||||||||||  ||||
Db   6391 CTCTGGGGCCAGAAGAAGAAGAACAGCAACATGACCTATGAGAAGCTGAGCCGAGCCAT 6449


RESULT  12
V58521/c
ID   V58521 standard; cDNA; 852 BP.
AC   V58521;
DT   08-DEC-1998  (first entry)
DE   Prostate tumour specific gene clone P18.
KW   Prostate tumour specific gene; human; prostate cancer; detection;
KW   therapy; ss.
OS   Homo sapiens.
PN   WO9837418-A2.
PD   27-AUG-1998.
PF   25-FEB-1998; U03690.
PR   09-FEB-1998; US-904809.
PR   25-FEB-1997; US-806596.
PR   01-AUG-1997; US-904809.
PA   (CORI-) CORIXA CORP.
PI   Dillon DC, Xu J;
DR   WPI; 98-480805/41.
PT   Novel human prostate specific tumour protein and fragments - useful
PT   for detecting and treating prostate cancers
PS   Claim 1; Page 55; 141pp; English.
CC   This sequence represents a human prostate tumour specific gene, and can
CC   be used in the method of the invention. The method is for detecting
CC   prostate cancer comprises contacting a biological sample with an agent
CC   able to bind an immunogenic portion of a prostate protein (such as
CC   encoded by this sequence). An antibody which binds to an immunogenic
CC   portion of the prostate protein, and the method can be used to detect,
CC   monitor progression of, or treat prostate cancers. The antibody may
CC   also be conjugated to a therapeutic agent for use in therapy of prostate
CC   cancers.
SQ   Sequence  852 BP;    174 A;    195 C;    206 G;    274 T;

  Query Match             3.2%;  Score 60.6;  DB 1;  Length 852;
  Best Local Similarity   56.2%;  Pred. No. 0.00043;
  Matches  114;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy   1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
           ||| | ||| | ||| ||   ||  | ||  | || ||  |||   |     | || ||
Db    388 CACTTATGGGAATTCATCCGCGACATCCTCTTGAACCCAGACAAGAACCCAGGATTAATA 329

Qy   1223 aggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccgg 1282
           |  |||    ||       ||||| ||||||   |   | | ||||     ||||| | |
Db    328 AAATGGGAAGACCGATCTGAGGGCGTCTTCAGGTTCTTGAAATCAGAGGCAGTGGCTCAG 269

Qy   1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatc 1342
           || ||||| |     ||||||     |  ||||| ||| || ||||| |||||  | ||
Db    268 CTATGGGGTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGAGCTATG 209

Qy   1343 cgccagtattacaagaagggcat 1365
           |  | || ||||| |  |  ||
Db    208 AGATATTACTACAAAAGAGAAAT 186


RESULT  13
V61167/c
ID   V61167 standard; cDNA; 852 BP.
AC   V61167;
DT   06-JAN-1999  (first entry)
DE   cDNA sequence of prostate tumour clone P18.
KW   Prostate; cancer; tumour; vaccine; immunogen; clone; ss.
OS   Homo sapiens.
PN   WO9837093-A2.
PD   27-AUG-1998.
PF   25-FEB-1998; U03492.
PR   09-FEB-1998; US-020956.
PR   25-FEB-1997; US-806099.
PR   01-AUG-1997; US-904804.
PA   (CORI-) CORIXA CORP.
PI   Dillon DC, Xu J;
DR   WPI; 98-609886/51.
PT   Polypeptides comprising immunogenic portions of prostate proteins -
PT   used in a vaccine for the treatment of prostate cancer
PS   Claim 3; Page 53; 130pp; English.
CC   The present sequence is a new DNA which encodes an immunogenic portion
CC   of a prostate tumour protein. The encoded immunogen, or the DNA itself,
CC   can be used as a vaccine for the treatment of prostate cancer. The DNA
```

```
CC   was identified by analysis of a subtracted cDNA library obtained by
CC   subtracting a prostate tumour cDNA expression library with a normal
CC   tissue cDNA library.
SQ   Sequence   852 BP;    174 A;    195 C;    206 G;   274 T;

  Query Match             3.2%;  Score 60.6;  DB 1;  Length 852;
  Best Local Similarity   56.2%;  Pred. No. 0.00043;
  Matches  114;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy    1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
           ||| | ||| | ||| ||    ||  | ||  | || ||  |||    |       | || ||
Db     388 CACTTATGGGAATTCATCCGCGACATCCTCTTGAACCCAGACAAGAACCCAGGATTAATA 329

Qy    1223 aggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccgg 1282
            |  |||    ||         ||||| ||||||   |     | | ||||     ||||| | |
Db     328 AAATGGGAAGACCGATCTGAGGGCGTCTTCAGGTTCTTGAAATCAGAGGCAGTGGCTCAG 269

Qy    1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatc 1342
           || |||||  |        ||||||       |  ||||| ||| || ||||| |||||   | ||
Db     268 CTATGGGGTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGAGCTATG 209

Qy    1343 cgccagtattacaagaagggcat 1365
             |   | || |||||  |   |  ||
Db     208 AGATATTACTACAAAAGAGAAAT 186

RESULT   14
V15529
ID   V15529 standard; cDNA; 2975 BP.
AC   V15529;
DT   22-MAY-1998  (first entry)
DE   Human lymphoid-specific transcription factor NERF-1 cDNA.
KW   ets-related; human; lymphoid-specific transcription factor; NERF-1;
KW   screening assay; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             122..1687
FT                   /*tag= a
FT                   /product= NERF-1
PN   US5721113-A.
PD   24-FEB-1998.
PF   03-JAN-1995; 368281.
PR   03-JAN-1995; US-368281.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Kunsch CA, Libermann TA, Oettgen JP;
DR   WPI; 98-168405/15.
DR   P-PSDB; W47237.
PT   NERF-1 and NERF-2 nucleic acids - encode ets-related human
PT   lymphoid-specific transcription factors
PS   Claim 1; Columns 17-22; 17pp; English.
CC   The present sequence encodes the ets-related lymphoid-specific
CC   transcription factor NERF-1, which can be used in screening assays
CC   for drugs that modulate NERF activity, and to treat patients having
CC   need of NERF-1. ets-related factors are involved in, e.g. cancer
CC   development, retrovirus replication, T-cell cytokine production,
CC   cell cycle effects, growth regulation and cell differentiation.
SQ   Sequence   2975 BP;    958 A;    575 C;    614 G;    828 T;

  Query Match             3.2%;  Score 60.4;  DB 1;  Length 2975;
  Best Local Similarity   53.9%;  Pred. No. 0.00056;
  Matches  124;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy    1151 gggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctat 1210
           ||  |   | | | |  ||||| |||| ||     ||  | || |    |       | |  | |
Db     518 GGAAACACAACCTATTTGTGGGAGTTTCTTTTAGATCTACTTCAAGATAAAAATACTTGT 577

Qy    1211 ggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcc 1270
             | | |   ||||  |||      | |   || || ||||| |||||   | | ||| |||
Db     578 CCCAGGTATATTAAATGGACTCAGAGAGAAAAAGGCATATTCAAGCTGGTGGATTCAAAG 637

Qy    1271 caggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctg 1330
             || |    ||| ||||| |    |  ||||||     || | ||||||||| || |   ||
Db     638 GCTGTCTCTAAGCTTTGGGGAAAGCATAAGAACAAACCAGACATGAACTATGAAACCATG 697

Qy    1331 agccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380
            | ||  |   |   |   | || ||| | | ||| ||   |    |||   ||
Db     698 GGACGAGCTTTGAGATACTACTACCAAAGGGGAATTCTTGCAAAGGTTGA 747

RESULT   15
V15530
ID   V15530 standard; cDNA; 3240 BP.
AC   V15530;
DT   22-MAY-1998  (first entry)
DE   Human lymphoid-specific transcription factor NERF-2 cDNA.
KW   ets-related; human; lymphoid-specific transcription factor; NERF-2;
KW   screening assay; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             207..1952
FT                   /*tag= a
FT                   /product= NERF-2
PN   US5721113-A.
PD   24-FEB-1998.
PF   03-JAN-1995; 368281.
PR   03-JAN-1995; US-368281.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Kunsch CA, Libermann TA, Oettgen JP;
DR   WPI; 98-168405/15.
DR   P-PSDB; W47238.
PT   NERF-1 and NERF-2 nucleic acids - encode ets-related human
PT   lymphoid-specific transcription factors
PS   Claim 2; Columns 23-28; 17pp; English.
CC   The present sequence encodes the ets-related lymphoid-specific
CC   transcription factor NERF-2, which can be used in screening assays
CC   for drugs that modulate NERF activity, and to treat patients having
CC   need of NERF-2. ets-related factors are involved in, e.g. cancer
CC   development, retrovirus replication, T-cell cytokine production,
CC   cell cycle effects, growth regulation and cell differentiation.
SQ   Sequence   3240 BP;   1070 A;    592 C;    662 G;    916 T;

  Query Match             3.2%;  Score 60.4;  DB 1;  Length 3240;
  Best Local Similarity   53.9%;  Pred. No. 0.00057;
  Matches  124;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy    1151 gggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctat 1210
           ||  |   | | | |  ||||| |||| ||     ||  | || |    |       | |  | |
Db     783 GGAAACACAACCTATTTGTGGGAGTTTCTTTTAGATCTACTTCAAGATAAAAATACTTGT 842

Qy    1211 ggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcc 1270
             | | |   ||||  |||      | |   || || ||||| |||||   | | ||| |||
Db     843 CCCAGGTATATTAAATGGACTCAGAGAGAAAAAGGCATATTCAAGCTGGTGGATTCAAAG 902

Qy    1271 caggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctg 1330
             || |    ||| ||||| |    |  ||||||     || | ||||||||| || |   ||
Db     903 GCTGTCTCTAAGCTTTGGGGAAAGCATAAGAACAAACCAGACATGAACTATGAAACCATG 962

Qy    1331 agccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380
            | ||  |   |   |   | || ||| | | ||| ||   |    |||   ||
Db     963 GGACGAGCTTTGAGATACTACTACCAAAGGGGAATTCTTGCAAAGGTTGA 1012

Search completed: November 20, 1999, 21:06:19
Job time: 1640 sec

                       GenCore version 4.5
              Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        November 20, 1999, 20:38:59 ; Search time 48.9 Seconds
```

(without alignments)
9690.466 Million cell updates/sec

Title: US-09-126-945-1
Perfect score: 1894
Sequence: 1 gtctgacttcctcccagcac.....ataaagatactagagaactg 1894

Scoring table: IDENTITY_NUCDX

Searched: 311585 seqs, 125096042 residues

Database : N_Geneseq_36:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	78.4	4.1	2667	1	T47198	Human ETS2 repress
2	78.2	4.1	1116	1	V38354	Transcription fact
3	78.2	4.1	1907	1	V38355	Transcription fact
4	78	4.1	2073	1	Q55149	Adenovirus E1A-F g
5	78	4.1	2064	1	T37087	E1AF matrix metall
6	76.6	4.0	1920	1	V40701	Human ELF3 protein
7	76.4	4.0	1800	1	Q91769	Coding sequence of
8	74.2	3.9	2938	1	Q50644	Human Hum-Fli-1 ge
9	74.2	3.9	1890	1	Q50662	Human Hum-Fli-1 ge
10	73.4	3.9	2410	1	V32688	Polyomavirus PEA3
11	68.6	3.6	7752	1	V38365	Mouse ESX transcri
c 12	60.6	3.2	852	1	V58521	Prostate tumour sp
c 13	60.6	3.2	852	1	V61167	cDNA sequence of p
14	60.4	3.2	2975	1	V15529	Human lymphoid-spe
15	60.4	3.2	3240	1	V15530	Human lymphoid-spe
16	60	3.2	2544	1	T47200	Mouse ETS2 repress
17	59.6	3.1	1604	1	V20465	Human c-ets oncoge
18	53.4	2.8	1447	1	X26551	DNA encoding chick
19	53.4	2.8	1528	1	X26552	DNA encoding chick
20	49	2.6	1364	1	V20472	Human c-spi oncoge
21	46.2	2.4	1913	1	Q15015	Nuclear factor C/E
22	45.8	2.4	15079	1	Q91580	S. clavuligerus cl
23	45.4	2.4	2386	1	Q63477	Micrococcal putres
24	45.2	2.4	15872	1	T68715	Streptomyces venez
25	45.2	2.4	43280	1	T80413	Tylactone synthase
26	44.6	2.4	2544	1	T67066	Human alpha3(IX) c
c 27	43	2.3	985	1	V44439	Mycobacterium tube
c 28	43	2.3	985	1	V64548	M. tuberculosis im
29	42.6	2.2	2338	1	Q14183	N.clavipes draglin
30	42.6	2.2	12001	1	Q76213	HSV L/ST region. H
31	42.6	2.2	2338	1	V23249	Nephila clavipes s
c 32	42.6	2.2	4626	1	X33449	Oryza sativa L. pi
33	42	2.2	2004	1	T85356	Nephila clavipes s
c 34	41.6	2.2	2568	1	T88419	Human glial cell l
c 35	41.6	2.2	1223	1	V00246	Human Ret ligand r
c 36	41.6	2.2	1682	1	V00247	Human Ret ligand r
c 37	41.6	2.2	2568	1	V99310	Glial cell-line de
c 38	41.6	2.2	1927	1	V99315	GDNFR-alpha cDNA c
c 39	41.6	2.2	1929	1	V99316	GDNFR-alpha cDNA c
40	41.4	2.2	53789	1	V21187	Amycolatopsis medi
c 41	41.2	2.2	738	1	X39714	Gastric cancer ass
42	41.2	2.2	927	1	X39716	Gastric cancer ass
43	41	2.2	1985	1	Q37480	Sequence encoding
44	40.6	2.1	23666	1	Q10190	Cephalosporin anti
45	40.6	2.1	44377	1	T78508	Platenolide syntha

ALIGNMENTS

RESULT 1

T47198
ID T47198 standard; cDNA; 2667 BP.
AC T47198;
DT 06-APR-1997 (first entry)
DE Human ETS2 repressor factor (ERF) cDNA.
KW ETS2 repressor factor; ERF; transcriptional repressor;
KW tumour suppressor; tumour; cancer; oncogene; gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 123. .1769
FT /*tag= a
FT polya_signal 2644. .2649
FT /*tag= b
PN WO9639517-A1.
PD 12-DEC-1996.
PF 04-JUN-1996; U10177.
PR 05-JUN-1995; US-469412.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
PI Athanasiou MA, Beal GJ, Blair DG, Fisher RJ, Mavrothalassitis GJ;
PI Sgouras D N;
DR WPI; 97-043139/04.
DR P-PSDB; W07700.
PT New DNA encoding ETS2 repressor factor - useful for reducing
PT tumourigenicity, esp. oncogene associated tumour cells
PS Claim 3; Page 63-65; 101pp; English.
CC A cDNA clone (T47198) codes for human ETS2 repressor factor (ERF)
CC (W07700). It was isolated from a K562 cDNA library using the H1
CC site of the ETS2 promoter as probe. A related clone (T47199)
CC coding for an alternatively spliced ERF (W07701) was also isolated.
CC The ERF gene, which maps to chromosome 19, q1.2-1.3, is the first
CC member of the ets family to be identified as a transcriptional
CC repressor in mammalian cells. It can be used to suppress or repress
CC transcription and to elucidate transcription process and regulation.
CC The ERF gene also has tumour suppresor activity and can be used to
CC reduce ets-dependent tumourigenicity associated with v-mos, c-met,
CC tpr-met, Ha-ras and gag-myb-ets oncogenes. The cDNA can be cloned
CC for expression of the ERF polypeptide in host cells.
SQ Sequence 2667 BP; 456 A; 876 C; 825 G; 510 T;

Query Match 4.1%; Score 78.4; DB 1; Length 2667;
Best Local Similarity 59.8%; Pred. No. 1.3e-07;
Matches 152; Conservative 0; Mismatches 96; Indels 6; Gaps 1;

```
Qy   1124 gacagcgaggtggactcatcatgctccgggcagcccatccacctgtggcagttcctcaag 1183
          | |  | ||   || || ||      |     ||  ||||| |||||||| ||  ||  |
Db    165 GCCTACAAGCCAGAGTCGTCCCCTGGCTCAAGGCAGATCCAGCTGTGGCACTTTATCCTG 224

Qy   1184 gagttgctactcaagccccacagctatggccgcttcattaggtggctcaacaaggagaag 1243
          ||| |||| |  |||     |  | | ||| | |     ||  |     || | | | |
Db    225 GAGCTGCTGCGGAAGGAGGAGTACCAGGGCGTCATTGCCTGGCAGGGGGACTACGGGGAA 284

Qy   1244 ggcatcttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaac 1303
            | || |||||      ||| | |   ||||||||||||||||||| | ||||||  |
Db    285 TTCGTCATCAAA------GACCCTGATGAGGTGGCCCGGCTGTGGGGCGTTCGCAAGTGC 338

Qy   1304 cgtcccgccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggc 1363
            |||  ||||| ||||||||||||||||  || | ||| | || || || ||| ||
Db    339 AAGCCCCAGATGAATTACGACAAGCTGAGCCGGGCCCTGCGCTATTACTATAACAAGCGC 398

Qy   1364 atcatccggaagcc 1377
          ||  | |  ||| |
Db    399 ATTCTGCACAAGAC 412
```

RESULT 2
V38354
ID V38354 standard; cDNA; 1116 BP.
AC V38354;
DT 09-NOV-1998 (first entry)
DE Transcription factor ESX gene open reading frame.
KW ESX; epithelial-restricted with serine box; transcription factor;
KW ETS; human; epithelial cancer; breast cancer; adenocarcinoma;

```
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   exon            1. .163
FT                   /*tag= a
FT                   /number= 2
FT   exon            164. .385
FT                   /*tag= b
FT                   /number= 3
FT   exon            386. .478
FT                   /*tag= c
FT                   /number= 3
FT   exon            479. .688
FT                   /*tag= d
FT                   /number= 5
FT   exon            689. .805
FT                   /*tag= e
FT                   /number= 6
FT   exon            806. .1001
FT                   /*tag= f
FT                   /number= 8
FT   exon            1002. .1111
FT                   /*tag= g
FT                   /number= 9
FT   misc_feature    1. .189
FT                   /*tag= h
FT                   /product= first variable region
FT                   /note= "Claim 28"
FT   misc_feature    190. .309
FT                   /*tag= i
FT                   /product= pointed region
FT   misc_feature    310. .561
FT                   /*tag= j
FT                   /product= second variable region
FT                   /note= "Claim 8"
FT   misc_feature    562. .714
FT                   /*tag= k
FT                   /product= serine-rich region
FT   misc_feature    715. .819
FT                   /*tag= l
FT                   /product= third variable region
FT                   /note= "Claim 28"
FT   misc_feature    820. .1062
FT                   /*tag= m
FT                   /product= Ets DNA binding domain
FT   misc_feature    1063. .1113
FT                   /*tag= n
FT                   /product= fourth variable region
FT                   /note= "Claim 28"
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60677-79.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
PT   useful for diagnosis and treatment of cancer
PS   Claim 3; Fig 1; 120pp; English.
CC   This is the open reading frame of the newly isolated ETX gene that
CC   codes for a novel transcription factor (see W60677) associated
CC   with the aetiology of cancers, including epithelial cancers.
CC   The gene is a member of the ETS family and is designated ETX for
CC   epithelial-restricted with serine box.  The gene is located at
CC   chromosome 1q32, a region amplified in 50% of early breast cancers.
CC   ESX is an early indicator of breast cancer, being expressed at the
CC   ductal carcinoma in situ stage.  Dysfunction of the ESX gene that
CC   results in increased expression is indicative of epithelial cancer
CC   (especially breast cancer but also gastric, ovarian and lung
CC   adenocarcinoma) and of unfavourable prognosis.  The ESX open
CC   reading frame was deduced from a cDNA clone of foetal liver-spleen
CC   origin that was identified in an expressed sequence tag database
CC   search using human Ets-2, and from a clone obtained by 5'RACE of
CC   placental polyA mRNA.  Antibodies, antisense nucleic acids or
CC   inactive ESX mutant proteins that inhibit activity of ESX can be
CC   used to inhibit growth and proliferation of neoplastic cells,
CC   particularly cancer.  ESX polypeptides, polynucleotides and
CC   antibodies are also used to detect ESX, particularly for diagnosis
CC   and monitoring of cancer.  Polynucleotides may also be used to
CC   detect gross abnormalities in chromosome 1, e.g. for in utero
CC   screening for predisposition to cancer, and to generate ESX protein
CC   in vivo from gene therapy vectors.  ESX protein, genes and cDNA
CC   can also be used to screen for specific modulators, especially
CC   potential anticancer agents.
SQ   Sequence  1116 BP;  255 A;  322 C;  335 G;  204 T;


  Query Match             4.1%;  Score 78.2;  DB 1;  Length 1116;
  Best Local Similarity   59.8%;  Pred. No. 1.2e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
           || |||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    816 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 875

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
           ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    876 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 935

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
           ||  || ||||||      || ||      |  ||||| |||||| ||||||||||  |
Db    936 CCAACTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 995

Qy   1339 catccgccagtattacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
           ||| |   ||| ||| | ||  |    |||| | | |  |       || ||| |||||
Db    996 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1052

Qy   1399 ctaccagttcg 1409
           |||| |||| |
Db   1053 CTACAAGTTTG 1063


RESULT   3
V38355
ID   V38355 standard; cDNA; 1907 BP.
AC   V38355;
DT   09-NOV-1998  (first entry)
DE   Transcription factor ESX cDNA (ORF and untranslated regions).
KW   ESX; epithelial-restricted with serine box; transcription factor;
KW   ETS; human; epithelial cancer; breast cancer; adenocarcinoma;
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   5'UTR           1. .95
FT                   /*tag= a
FT   CDS             96. .1206
FT                   /*tag= b
FT                   /note= "Claim 3"
FT   3'UTR           1207. .1907
FT                   /*tag= c
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60677.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
```

```
PT   useful for diagnosis and treatment of cancer
PS   Disclosure; Page 83; 120pp; English.
CC   This cDNA sequence comprises the open reading frame (see also
CC   V38354) and 5' and 3' untranslated regions of the newly isolated
CC   ETX gene that codes for a novel human transcription factor (see
CC   W60677) associated with the aetiology of cancers. The ETX gene
CC   is a member of the ETS family and is designated ETX for epithelial
CC   restricted with serine box. The gene is located at chromosome
CC   1q32, a region amplified in 50% of early breast cancers. ESX is an
CC   early indicator of breast cancer, being expressed at the ductal
CC   carcinoma in situ stage. Dysfunction of the ESX gene that results
CC   in increased expression is indicative of epithelial cancer
CC   (especially breast cancer but also gastric, ovarian and lung
CC   adenocarcinoma) and of unfavourable prognosis. The ESX open
CC   reading frame was deduced from a cDNA clone of foetal liver-spleen
CC   origin that was identified in an expressed sequence tag database
CC   search using human Ets-2, and from a clone obtained by 5'RACE of
CC   placental polyA mRNA. Antibodies, antisense nucleic acids or
CC   inactive ESX mutant proteins that inhibit activity of ESX can be
CC   used to inhibit growth and proliferation of neoplastic cells,
CC   particularly cancer. ESX polynucleotides, polypeptides and
CC   antibodies are also used to detect ESX, particularly for diagnosis
CC   and monitoring of cancer. ESX polynucleotides may also be used to
CC   detect gross abnormalities in chromosome 1, e.g. for in utero
CC   screening for predisposition to cancer, and to generate ESX protein
CC   in vivo from gene therapy vectors. ESX protein, genes and cDNA can
CC   also be used to screen for specific modulators, especially
CC   potential anticancer agents.
SQ   Sequence   1907 BP;    427 A;    557 C;    533 G;    390 T;

  Query Match             4.1%;  Score 78.2;  DB 1;  Length 1907;
  Best Local Similarity   59.8%;  Pred. No. 1.3e-07;
  Matches  150;  Conservative    0;  Mismatches   98;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
           || ||||||||||| ||||| ||  |||  | ||  || | ||  |   | | |   || |
Db    911 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 970

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||   ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    971 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 1030

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
          ||  || ||||||      || ||       |  ||||| |||||| ||||||||||  |
Db   1031 CCAACTATGGGGCCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 1090

Qy   1339 catccgccagtattacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||| ||| | ||  |    |||| | | |         || ||| ||||
Db   1091 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1147

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1148 CTACAAGTTTG 1158


RESULT   4
Q55149
ID   Q55149 standard; cDNA to mRNA; 2073 BP.
AC   Q55149;
DT   11-JUL-1994  (first entry)
DE   Adenovirus E1A-F gene.
KW   Adenovirus; cancer; ets oncogene; HeLa cell; enhancer core sequence;
KW   methylation; ds.
OS   Human adenovirus.
FH   Key             Location/Qualifiers
FT   cds             844. .1311
FT                   /*tag= a
FT                   /note= "Claimed sequence"
PN   J05328975-A.
PD   14-DEC-1993.
PF   02-JUN-1992; 165453.
PR   02-JUN-1992; JP-165453.
PA   (TAKI ) TAKARA SHUZO CO LTD.
DR   WPI; 94-021923/03.
DR   P-PSDB; R45451.
PT   Novel E1A-F gene - for production of adenovirus E1A-F and cancer
PT   research
PS   Claim 1; Page 6; 7pp; Japanese.
CC   The adenovirus E1A-F gene contains a 473bp open reading frame. The
CC   clone comprising the coding sequence was isolated by screening
CC   a HeLa cell cDNA library.
SQ   Sequence   2073 BP;    458 A;    635 C;    561 G;    418 T;

  Query Match             4.1%;  Score 78;  DB 1;  Length 2073;
  Best Local Similarity   76.2%;  Pred. No. 1.5e-07;
  Matches  96;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
           |||||  |    ||  | |   |||| ||| |||| ||||||||||  ||||||||  ||
Db    961 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1020

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
          ||||||||  ||||||||||||||||||||  ||||  | |||||  |||| |||||||||
Db   1021 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1080

Qy   1370 cggaag 1375
          |  ||||
Db   1081 CAGAAG 1086


RESULT   5
T37087
ID   T37087 standard; cDNA to mRNA; 2064 BP.
AC   T37087;
DT   25-APR-1997  (first entry)
DE   E1AF matrix metalloproteinase regulator, cDNA.
KW   E1AF; matrix; metalloproteinase; regulator; infiltration; cancer;
KW   metastasis; cell; control; antisense; decoy; DNA binding region;
KW   target DNA; ribosome; induction; diagnosis; detection; treatment;
KW   mammary cancer; fibrosarcoma; osteosarcoma; lung cancer; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .1389
FT                   /*tag= a
PN   WO9624379-A1.
PD   15-AUG-1996.
PF   09-JAN-1996; J00016.
PR   08-FEB-1995; JP-020173.
PA   (TAKI ) TAKARA SHUZO CO LTD.
PI   Fujinaga K,  Higashino F,  Yoshida K;
DR   WPI; 96-384227/38.
DR   P-PSDB; W00167.
PT   Control of cancer cell infiltration by E1AF gene expression
PT   regulation  - also diagnosis of cancer by detection of E1AF gene
PT   expression products
PS   Example 5; Pages 38-42; 65pp; Japanese.
CC   The present sequence encodes the E1AF protein, which is a matrix
CC   metalloproteinase regulator. The infiltration and metastasis of
CC   cancer cells can be controlled by regulating the expression and
CC   expression products of the E1AF gene. This may be accomplished by
CC   inducing antisense DNA or RNA for the E1AF gene, a decoy gene
CC   expressing the DNA binding region of the E1AF protein, the target
CC   DNA for the DNA binding region of the E1AF protein or ribosomes
CC   corresponding to the E1AF gene mRNA. Cancer can be diagnosed by
CC   detecting E1AF gene expression products, e.g. E1AF protein or mRNA.
CC   These methods may be used in the treatment and diagnosis of cancer,
CC   e.g. mammary cancer, fibrosarcoma, osteosarcoma, lung cancer, etc. .
SQ   Sequence   2064 BP;    427 A;    648 C;    569 G;    420 T;

  Query Match             4.1%;  Score 78;  DB 1;  Length 2064;
  Best Local Similarity   76.2%;  Pred. No. 1.5e-07;
  Matches  96;  Conservative    0;  Mismatches   30;  Indels    0;  Gaps    0;
```

```
Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          ||||||  |    ||  | |   |||| ||| |||| ||||||||||  |||||||| ||
Db   1039 TTCAAGCTCATTGAGCCTGAGGAGGTCGCCAGGCTCTGGGGCATCCAGAAGAACCGGCCA 1098

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
          |||||||| ||||||||||||||||||||  ||||  | |||||  |||| ||||||||
Db   1099 GCCATGAATTACGACAAGCTGAGCCGCTCGCTCCGATACTATTATGAGAAAGGCATCATG 1158

Qy   1370 cggaag 1375
          | ||||
Db   1159 CAGAAG 1164


RESULT   6
V40701
ID   V40701 standard; DNA; 1920 BP.
AC   V40701;
DT   21-SEP-1998  (first entry)
DE   Human ELF3 protein coding sequence.
KW   ELF3; human; cancer; epithelial cancer; breast cancer; prostate cancer;
KW   diagnosis; therapy; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             115. .1230
FT                   /*tag= a
FT                   /product= ELF3
PN   EP-839908-A2.
PD   06-MAY-1998.
PF   23-OCT-1997; 308450.
PR   15-NOV-1996; US-746789.
PR   31-OCT-1996; US-028791.
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (MONU ) UNIV MONASH.
PI   Debouck C, Kola I, Tymms MJ;
DR   WPI; 98-242674/22.
DR   P-PSDB; W57834.
PT   New isolated polynucleotide and e.g. DNA, agonists and antagonists -
PT   used to diagnose and treat cancer, e.g. prostate, breast, lung or
PT   epithelial tumours
PS   Claim 4; Fig 1A; 34pp; English.
CC   This sequence encodes the human ELF3 protein of the invention. ELF3 is
CC   used to treat patients requiring ELF3 by providing DNA encoding the
CC   polypeptide and expressing the polypeptide in vivo. ELF3 is used to
CC   diagnose a disease or a susceptibility to a disease related to its
CC   expression by determining a mutation in the nucleic acid sequence
CC   encoding it. The presence ELF3 in a sample derived from a host can be
CC   detected. Antagonists of ELF3 are used to treat patients requiring
CC   inhibition of ELF3 polypeptide. ELF3 can be used in the early diagnosis
CC   and treatment of cancer, particularly epithelial cancer such as breast
CC   and prostate cancer.
SQ   Sequence   1920 BP;    422 A;    561 C;    550 G;    387 T;


  Query Match             4.0%;  Score 76.6;  DB 1;  Length 1920;
  Best Local Similarity   59.4%;  Pred. No. 2.8e-07;
  Matches  149;  Conservative    0;  Mismatches   99;  Indels    3;  Gaps    1;

Qy   1159 catccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgctt 1218
          || |||||||||| ||||| ||  |||  | ||  || | ||  |    | | |    || |
Db    930 CACCCACCTGTGGGAGTTCATCCGGGACATCCTCATCCACCCGGAGCTCAACGAGGGCCT 989

Qy   1219 cattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggc 1278
          ||| | ||||    ||   | |  | ||| |||||||  |   |  ||| |     |||||
Db    990 CATGAAGTGGGAGAATCGGCATGAAGGCGTCTTCAAGTTCCTGCGCTCCGAGGCTGTGGC 1049

Qy   1279 ccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctc 1338
          ||  || |||||       || ||     |  ||||| |||||| ||||||||||  |
Db   1050 CCAACTATGGGGGCAAAAGAAAAAGAACAGCAACATGACCTACGAGAAGCTGAGCCGGGC 1109

Qy   1339 catccgccagtattacaagaagggcatcatccggaagccagacatctcccagcgcctcgt 1398
          |||  |   ||| ||| | ||  |    |||| | | |  |        || ||| |||||
Db   1110 CATGAG---GTACTACTACAAACGGGAGATCCTGGAACGGGTGGATGGCCGGCGACTCGT 1166

Qy   1399 ctaccagttcg 1409
          |||| |||| |
Db   1167 CTACAAGTTTG 1177


RESULT   7
Q91769
ID   Q91769 standard; cDNA to mRNA; 1800 BP.
AC   Q91769;
DT   09-FEB-1996  (first entry)
DE   Coding sequence of PEA3-beta, an ETS transcription factor.
KW   Transcription factor; probe; reverse transcription; PCR; primer;
KW   expression vector; E.coli; COS cell; ras; cancer cell multiplication;
KW   polyoma virus; transformation; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             1. .1563
FT                   /*tag= a
FT                   /product= PEA3-beta transcription factor
FT   cds             250. .252
FT                   /*tag= b
FT                   /transl_except= seq: CTC  a.a.:Pro
FT   cds             376. .378
FT                   /*tag= c
FT                   /transl_except= seq: AAC  a.a.: Lys
FT   cds             1498. .1500
FT                   /*tag= d
FT                   /transl_except= seq: ACC  a.a.: Ser
PN   J07145197-A.
PD   06-JUN-1995.
PF   25-NOV-1993; 295393.
PR   25-NOV-1993; JP-295393.
PA   (EISA ) EISAI CO LTD.
PA   (HIRA/) HIRANO T.
DR   WPI; 95-237197/31.
DR   P-PSDB; R78185.
PT   ETS transcription factor activated by ras - may be used in the study
PT   of cancer cell proliferation and the proliferation of the polyoma
PT   virus
PS   Claim 2; Page 7-9; 9pp; Japanese.
CC   The nucleotide sequence of the novel ETS transcription factor family
CC   member - PEA3-beta.  The gene was isolated from a HepG2 cell line cDNA
CC   library.  The probe for the screening was prepared by reverse
CC   transcription on HepG2 mRNA followed by PCR using primers Q91770-1,
CC   to produce a probe of 170-200 bp.  The screening isolated the full
CC   length sequence of the transcription factor.  The gene was inserted into
CC   the expression vectors pBluescript KS and pcDV1 for expression of the
CC   protein in E.coli and COS7 cells, respectively.  The ETS transcription
CC   factor has specificity for and is activated by ras. It is useful as a
CC   reagent in studies for the elucidation of the mechanism of cancer cell
CC   multiplication or polyoma virus transformation of cells.
SQ   Sequence   1800 BP;    437 A;    522 C;    435 G;    405 T;


  Query Match             4.0%;  Score 76.4;  DB 1;  Length 1800;
  Best Local Similarity   75.4%;  Pred. No. 3e-07;
  Matches   95;  Conservative    0;  Mismatches   31;  Indels    0;  Gaps    0;

Qy   1250 ttcaaaattgaggactcagcccaggtggcccggctgtggggcatccgcaagaaccgtccc 1309
          |||||  |    ||  | |   |||| || |||| ||||||||||  |||||||| ||
Db   1186 TTCAAGCTGATAGAACCGGAAGAGGTTGCTCGGCGCTGGGGCATCCAGAAGAACCGGCCA 1245

Qy   1310 gccatgaactacgacaagctgagccgctccatccgccagtattacaagaagggcatcatc 1369
          |||||||||| ||||||||||||||||||  ||||| | || ||  | ||||||||||
Db   1246 GCCATGAACTATGACAAGCTGAGCCGCTCTCTCCGCTATTACTATGAAAAGGGCATCATG 1305

Qy   1370 cggaag 1375
          | ||||
Db   1306 CAGAAG 1311


RESULT   8
Q50644
```

```
ID   Q50644 standard; cDNA; 2938 BP.
AC   Q50644;
DT   26-MAY-1994  (first entry)
DE   Human Hum-Fli-1 gene clone BM025.
KW   chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
KW   Ews gene; malignant melanoma; hum-fli-1;
KW   primitive peripheral neuroectodermal tumour; human chromosome 11;
KW   human chromosome 22; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   cds             143. .1501
FT                   /*tag= a
FT                   /product= HUM-FLI-1
FT   polya_signal    2908. .2913
FT                   /*tag= b
PN   WO9323549-A.
PD   25-NOV-1993.
PF   19-MAY-1993; F00494.
PR   20-MAY-1992; FR-006123.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PI   Aurias A,  Delattre O,  Desmaze C,  Melot T,  Peter M;
PI   Plougastel B, Thomas G,  Zucman J;
DR   WPI; 93-386580/48.
DR   P-PSDB; R44556.
PT   New nucleic acid of EWS gene and its hybrid(s) - contg. gene
PT   sequence involved in chromosomal trans-location, also derived
PT   mRNA, probes, fusion proteins etc., for diagnosis and treatment
PT   of Ewing sarcoma and melanoma
PS   Disclosure; Fig 7; 123pp; French.
CC   The probe 11RR1 was used to screen a human marrow cDNA library
CC   (Clontech cat.# HL1058). The clone BM025 was identified and
CC   sequenced. It represents the entire coding region together with
CC   5'- and 3'-UTRs of the Hum-Fli-1 gene.
SQ   Sequence   2938 BP;    847 A;    692 C;    675 G;    724 T;


  Query Match             3.9%;  Score 74.2;  DB 1;  Length 2938;
  Best Local Similarity   62.9%;  Pred. No. 9e-07;
  Matches  134;  Conservative    0;  Mismatches   73;  Indels    6;  Gaps    1;

Qy   1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
           ||  ||||| |||||||| ||||||  |||| ||||   | |   |  || |    || |
Db    979 GCAGATCCAGCTGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTG 1038

Qy   1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggt 1275
          |||   ||  |   | ||| | | ||      |||||||  |||| | |   ||||
Db   1039 TATCACCTGGGAGGGGACCAACGGGGAG------TTCAAAATGACGGACCCCGATGAGGT 1092

Qy   1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccg 1335
          |||| |||   ||||||   || || | |   |||  |||||| ||||||||||||||||
Db   1093 GGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAAGCTGAGCCG 1152

Qy   1336 ctccatccgccagtattacaagaagggcatcat 1368
           || ||||  | || ||  | ||   ||| ||
Db   1153 GGCCCTCCGTTATTACTATGATAAAAACATTAT 1185


RESULT   9
Q50662
ID   Q50662 standard; DNA; 1890 BP.
AC   Q50662;
DT   26-MAY-1994  (first entry)
DE   Human Hum-Fli-1 gene (genomic DNA).
KW   chromosomal translocation; chimeric; chimaeric; Ewing sarcoma;
KW   Ews gene; malignant melanoma; hum-fli-1;
KW   primitive peripheral neuroectodermal tumour; human chromosome 11;
KW   human chromosome 22; ss.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   exon            1. .39
FT                   /*tag= a
FT                   /number= 1
FT                   /codon_start= 22. .24
FT   intron          40. .99
FT                   /*tag= b
FT                   /number= 1
FT   exon            100. .311
FT                   /*tag= c
FT                   /number= 2
FT   intron          312. .371
FT                   /*tag= d
FT                   /number= 2
FT   exon            372. .526
FT                   /*tag= e
FT                   /number= 3
FT   intron          527. .586
FT                   /*tag= f
FT                   /number= 3
FT   exon            587. .790
FT                   /*tag= g
FT                   /number= 4
FT   intron          791. .850
FT                   /*tag= h
FT                   /number= 4
FT   exon            851. .916
FT                   /*tag= i
FT                   /number= 5
FT   intron          917. .976
FT                   /*tag= j
FT                   /number= 5
FT   exon            977. .1042
FT                   /*tag= k
FT                   /number= 6
FT   intron          1043. .1102
FT                   /*tag= l
FT                   /number= 6
FT   exon            1103. .1162
FT                   /*tag= m
FT                   /number= 7
FT   intron          1163. .1222
FT                   /*tag= n
FT                   /number= 7
FT   exon            1223. .1270
FT                   /*tag= o
FT                   /number= 8
FT   intron          1271. .1330
FT                   /*tag= p
FT                   /number= 8
FT   exon            1331. .1890
FT                   /*tag= q
FT                   /number= 9
PN   WO9323549-A.
PD   25-NOV-1993.
PF   19-MAY-1993; F00494.
PR   20-MAY-1992; FR-006123.
PA   (CNRS ) CNRS CENT NAT RECH SCI.
PI   Aurias A,  Delattre O,  Desmaze C,  Melot T,  Peter M;
PI   Plougastel B, Thomas G,  Zucman J;
DR   WPI; 93-386580/48.
DR   P-PSDB; R44556.
PT   New nucleic acid of EWS gene and its hybrid(s) - contg. gene
PT   sequence involved in chromosomal trans-location, also derived
PT   mRNA, probes, fusion proteins etc., for diagnosis and treatment
PT   of Ewing sarcoma and melanoma
PS   Disclosure; Fig 7 and Fig 13; 123pp; French.
CC   The intron-exon junctions of the human Ews gene and the Hum-Fli-1
CC   gene have been sequenced (see Q50646 and Q50662, respectively).
CC   The different fusion products which could be formed by fusing
CC   exons from the two genes, as happens after specific chromosomal
CC   translocations, can be predicted (see Q50671-Q50678). The sequences
CC   at fusion junctions of other observed translocations are given in
CC   Q50679-Q50683). NOTE: the 1890 residue sequence Q50662 does not
CC   appear in the specification; it is a combination of the Hum-Fli-1
CC   cDNA sequence (Fig 7, Q50644) and the individual intron sequences
CC   with their intron-exon junctions (Fig 13).
SQ   Sequence   1890 BP;    457 A;    561 C;    458 G;    414 T;
```

```
  Query Match             3.9%;  Score 74.2;  DB 1;  Length 1890;
  Best Local Similarity   62.9%;  Pred. No. 8.5e-07;
  Matches  134;  Conservative    0;  Mismatches   73;  Indels    6;  Gaps    1;

Qy    1156 gcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccg 1215
           ||  ||||| |||||||| ||||||  |||| ||||   | |    |  || |    || |
Db    1338 GCAGATCCAGCTGTGGCAATTCCTCCTGGAGCTGCTCTCCGACAGCGCCAACGCCAGCTG 1397

Qy    1216 cttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggt 1275
             |||   ||  |   | ||| | | ||      ||||||||   |||| | |   ||||
Db    1398 TATCACCTGGGAGGGGACCAACGGGGAG------TTCAAAATGACGGACCCCGATGAGGT 1451

Qy    1276 ggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccg 1335
           |||| |||  ||||||   || || | |   |||  |||||| ||||||||||||||||||
Db    1452 GGCCAGGCGCTGGGGCGAGCGGAAAAGCAAGCCCAACATGAATTACGACAAGCTGAGCCG 1511

Qy    1336 ctccatccgccagtattacaagaagggcatcat 1368
             || ||||  | || ||  | ||   ||| ||
Db    1512 GGCCCTCCGTTATTACTATGATAAAAACATTAT 1544


RESULT  10
V32688
ID   V32688 standard; cDNA; 2410 BP.
AC   V32688;
DT   20-OCT-1998  (first entry)
DE   Polyomavirus PEA3 cDNA.
KW   Polyomavirus enhancer activator; PEA3; tumour; suppressor; inhibitor;
KW   transformation; HER-2; neu promoter; metastasis; cancer; ss.
OS   Polyomavirus.
FH   Key             Location/Qualifiers
FT   CDS             126. .1793
FT                   /*tag=  a
FT                   /product= PEA3
FT                   /note= "polyomavirus enhancer activator"
PN   WO9830585-A2.
PD   16-JUL-1998.
PF   12-JAN-1998; U00880.
PR   10-JAN-1997; US-780835.
PA   (TEXA ) UNIV TEXAS SYSTEM.
PI   Hung M, Xing X;
DR   WPI; 98-399061/34.
DR   P-PSDB; W49010.
PT   Method for repressing transformation of cells - by contacting cell
PT   with polyoma-virus enhancer activator 3, useful for, e.g. treating
PT   or preventing cancer, tumourigenesis and metastasis
PS   Disclosure; Page 69-70; 83pp; English.
CC   This sequence encodes a polyomavirus enhancer activator, PEA3. This
CC   sequence is used in a method for repressing transformation in a cell
CC   which involves contacting the cell with PEA3 to inhibit a transformed
CC   phenotype. This sequence can also be used in a method to suppress the
CC   growth of a tumour in a mammal comprising introducing to the mammal a
CC   PEA3-encoding nucleic acid where the expression of PEA3 in the mammal
CC   results in a decrease in the growth rate of the tumour. PEA3 can regulate
CC   the HER-2/neu promoter by transcriptional repression and acts as a tumour
CC   repressor. The PEA3 can be used for reducing a transforming,
CC   tumourogenic or metastatic potential of a cell. It can be used for the
CC   prevention and treatment of such transformation-driven events as cancer,
CC   tumourigenesis and metastasis.
SQ   Sequence  2410 BP;    521 A;    744 C;    645 G;    500 T;


  Query Match             3.9%;  Score 73.4;  DB 1;  Length 2410;
  Best Local Similarity   59.9%;  Pred. No. 1.3e-06;
  Matches  160;  Conservative    0;  Mismatches  101;  Indels    6;  Gaps    2;

Qy    1109 gaggagagctggaccgacagcgaggtggactcatcatgctccgggcagcccatccacctg 1168
           || | ||   |  |   | | |||| |    | |       | ||    || | || |||
Db    1308 GAAGGGATTGGAGCTTTCCGGGAGGGGCCACCCTACCAGCGCCGGGGTGCCTTACAACTG 1367

Qy    1169 tggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcattaggtgg 1228
           |||||||| ||   || |   || ||  |  ||  |   |||  |  ||||||   |||
Db    1368 TGGCAGTTTCT---GGTGGCCCTGCTGGATGACCCAACAAATGCTCATTTCATTGCTTGG 1424

Qy    1229 ctcaacaaggagaagggcatcttcaaaattgaggactcagcccaggtggcccggctgtgg 1288
             |  ||  | ||     || ||| |    ||  | |   |||| ||| |||| |||
Db    1425 ACAGGCCGGGGAATGG---AGTTTAAACTAATTGAACCTGAAGAGGTTGCCAGGCTCTGG 1481

Qy    1289 ggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatccgccag 1348
           || ||||  |||||||| || |||||||| || ||||||||||||||||  | ||  |
Db    1482 GGTATCCAGAAGAACCGGCCAGCCATGAATTATGACAAGCTGAGCCGCTCGCTGCGATAC 1541

Qy    1349 tattacaagaagggcatcatccggaag 1375
           |||||  |||| |||||||| | ||||
Db    1542 TATTATGAGAAAGGCATCATGCAGAAG 1568


RESULT  11
V38365
ID   V38365 standard; cDNA; 7752 BP.
AC   V38365;
DT   09-NOV-1998  (first entry)
DE   Mouse ESX transcription factor gene.
KW   ESX; epithelial-restricted with serine box; transcription factor;
KW   ETS; mouse; epithelial cancer; breast cancer; adenocarcinoma;
KW   diagnosis; therapy; ss.
OS   Mus sp.
FH   Key             Location/Qualifiers
FT   promoter        2553. .2889
FT                   /*tag=  a
FT   CDS             3604. .7115
FT                   /*tag=  b
FT                   /note= "contains introns"
FT   exon            2900. .3035
FT                   /*tag=  c
FT                   /number= 1
FT   intron          3036. .3606
FT                   /*tag=  d
FT                   /number= 1
FT   exon            3604. .3764
FT                   /*tag=  e
FT                   /number= 2
FT   intron          3765. .4148
FT                   /*tag=  f
FT                   /number= 2
FT   exon            4149. .4373
FT                   /*tag=  g
FT                   /number= 3
FT   intron          4374. .4503
FT                   /*tag=  h
FT                   /number= 3
FT   exon            4504. .4594
FT                   /*tag=  i
FT                   /number= 4
FT   intron          4595. .4782
FT                   /*tag=  j
FT                   /number= 4
FT   exon            4783. .4905
FT                   /*tag=  k
FT                   /number= 5
FT   intron          4906. .5052
FT                   /*tag=  l
FT                   /number= 5
FT   exon            5053. .5142
FT                   /*tag=  m
FT                   /number= 6
FT   intron          5143. .5284
FT                   /*tag=  n
FT                   /number= 6
FT   exon            5285. .5402
FT                   /*tag=  o
FT                   /number= 7
FT   intron          5403. .6255
FT                   /*tag=  p
```

```
FT                   /number= 7
FT   exon            6256. .6450
FT                   /*tag=  q
FT                   /number= 8
FT   intron          6451. .6998
FT                   /*tag=  r
FT                   /number= 8
FT   exon            6999. .7115
FT                   /*tag=  s
FT                   /number= 9
PN   WO9823782-A2.
PD   04-JUN-1998.
PF   26-NOV-1997; U21865.
PR   25-NOV-1997; US-031504.
PR   27-NOV-1996; US-031504.
PA   (REGC ) UNIV CALIFORNIA.
PI   Benz CC, Chang C, Scott GK;
DR   WPI; 98-322755/28.
DR   P-PSDB; W60680.
PT   Nucleic acid encoding the ESX transcription activator -
PT   over-expressed in epithelial, especially breast, cancers, also
PT   related polypeptide(s), antibodies, vectors and transformed cells,
PT   useful for diagnosis and treatment of cancer
PS   Claim 18; Page 85-90; 120pp; English.
CC   This is the murine mETX gene that contains about 2.9 kb of
CC   promoter upstream of about 4.9 kb of DNA incorporating at least
CC   9 exons that specify a transcript of about 2 kb, with exons 2-9
CC   encoding the 371-amino acid mESX protein (see W60680).  The gene
CC   is a member of the ETS family and is designated ETX for
CC   epithelial-restricted with serine box.  In humans, a dysfunction of
CC   the ETX gene (see V38354) that results in increased expression is
CC   indicative of epithelial cancer (especially breast cancer but
CC   also gastric, ovarian and lung adenocarcinoma) and of unfavourable
CC   prognosis.  Antibodies, antisense nucleic acids or inactive ESX
CC   mutant proteins that inhibit activity of human ESX can be used to
CC   inhibit growth and proliferation of neoplastic cells, particularly
CC   cancer.  ESX polypeptides, polynucleotides and antibodies are also
CC   used to detect ESX, particularly for diagnosis and monitoring of
CC   cancer.  They can also be used to screen for specific modulators,
CC   especially potential anticancer agents.  Transgenic animals are
CC   also provided.
SQ   Sequence   7752 BP;    1768 A;    2050 C;    2019 G;    1914 T;

  Query Match             3.6%;  Score 68.6;  DB 1;  Length 7752;
  Best Local Similarity   61.5%;  Pred. No. 1.4e-05;
  Matches  110;  Conservative    0;  Mismatches   69;  Indels    0;  Gaps    0;

Qy   1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
           |||||||||| ||||  ||   ||  | ||| || |  ||| |  | | |    || ||||
Db   6271 CACCTGTGGGAGTTTATCCGAGACATCCTAATCCACCCCGAGCTCAACGAAGGCCTCATG 6330

Qy   1223 aggtggctcaacaaggagaaggggcatcttcaaaattgaggactcagcccaggtggccccgg 1282
          | ||||   |||  | |  ||||  | |||||  ||       |||||      ||||||| |
Db   6331 AAGTGGGAGAACCGGCACGAGGGTGTGTTCAAGTTTCTTCGCTCAGAGGCCGTGGCCCAA 6390

Qy   1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccat 1341
           || ||||||      |||||       |  ||||| ||| || ||||||||||  ||||
Db   6391 CTCTGGGGCCAGAAGAAGAAGAACAGCAACATGACCTATGAGAAGCTGAGCCGAGCCAT 6449


RESULT  12
V58521/c
ID   V58521 standard; cDNA; 852 BP.
AC   V58521;
DT   08-DEC-1998  (first entry)
DE   Prostate tumour specific gene clone P18.
KW   Prostate tumour specific gene; human; prostate cancer; detection;
KW   therapy; ss.
OS   Homo sapiens.
PN   WO9837418-A2.
PD   27-AUG-1998.
PF   25-FEB-1998; U03690.
PR   09-FEB-1998; US-904809.
PR   25-FEB-1997; US-806596.
PR   01-AUG-1997; US-904809.
PA   (CORI-) CORIXA CORP.
PI   Dillon DC, Xu J;
DR   WPI; 98-480805/41.
PT   Novel human prostate specific tumour protein and fragments - useful
PT   for detecting and treating prostate cancers
PS   Claim 1; Page 55; 141pp; English.
CC   This sequence represents a human prostate tumour specific gene, and can
CC   be used in the method of the invention. The method is for detecting
CC   prostate cancer comprises contacting a biological sample with an agent
CC   able to bind an immunogenic portion of a prostate protein (such as
CC   encoded by this sequence). An antibody which binds to an immunogenic
CC   portion of the prostate protein, and the method can be used to detect,
CC   monitor progression of, or treat prostate cancers. The antibody may
CC   also be conjugated to a therapeutic agent for use in therapy of prostate
CC   cancers.
SQ   Sequence   852 BP;    174 A;    195 C;    206 G;    274 T;

  Query Match             3.2%;  Score 60.6;  DB 1;  Length 852;
  Best Local Similarity   56.2%;  Pred. No. 0.00043;
  Matches  114;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy   1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
           ||| | ||| | ||| ||   ||  | ||  | || ||  |||   |     | || ||
Db    388 CACTTATGGGAATTCATCCGCGACATCCTCTTGAACCCAGACAAGAACCCAGGATTAATA 329

Qy   1223 aggtggctcaacaaggagaaggggcatcttcaaaattgaggactcagcccaggtggccccgg 1282
          |  |||   .||       ||||| ||||||    |   | | ||||      ||||| | |
Db    328 AAATGGGAAGACCGATCTGAGGGCGTCTTCAGGTTCTTGAAATCAGAGGCAGTGGCTCAG 269

Qy   1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatc 1342
          || ||||| |     ||||||     |  ||||| ||| || ||||| |||||  | ||
Db    268 CTATGGGGTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGAGCTATG 209

Qy   1343 cgccagtattacaagaagggcat 1365
           |  | || ||||| |  |  ||
Db    208 AGATATTACTACAAAAGAGAAAT 186


RESULT  13
V61167/c
ID   V61167 standard; cDNA; 852 BP.
AC   V61167;
DT   06-JAN-1999  (first entry)
DE   cDNA sequence of prostate tumour clone P18.
KW   Prostate; cancer; tumour; vaccine; immunogen; clone; ss.
OS   Homo sapiens.
PN   WO9837093-A2.
PD   27-AUG-1998.
PF   25-FEB-1998; U03492.
PR   09-FEB-1998; US-020956.
PR   25-FEB-1997; US-806099.
PR   01-AUG-1997; US-904804.
PA   (CORI-) CORIXA CORP.
PI   Dillon DC, Xu J;
DR   WPI; 98-609886/51.
PT   Polypeptides comprising immunogenic portions of prostate proteins -
PT   used in a vaccine for the treatment of prostate cancer
PS   Claim 3; Page 53; 130pp; English.
CC   The present sequence is a new DNA which encodes an immunogenic portion
CC   of a prostate tumour protein. The encoded immunogen, or the DNA itself,
CC   can be used as a vaccine for the treatment of prostate cancer. The DNA
CC   was identified by analysis of a subtracted cDNA library obtained by
CC   subtracting a prostate tumour cDNA expression library with a normal
CC   tissue cDNA library.
SQ   Sequence   852 BP;    174 A;    195 C;    206 G;    274 T;

  Query Match             3.2%;  Score 60.6;  DB 1;  Length 852;
  Best Local Similarity   56.2%;  Pred. No. 0.00043;
```

```
  Matches  114;  Conservative    0;  Mismatches   89;  Indels    0;  Gaps    0;

Qy   1163 cacctgtggcagttcctcaaggagttgctactcaagccccacagctatggccgcttcatt 1222
          ||| | ||| | ||| ||   ||  | ||  | || ||  |||    |      | || ||
Db    388 CACTTATGGGAATTCATCCGCGACATCCTCTTGAACCCAGACAAGAACCCAGGATTAATA 329

Qy   1223 aggtggctcaacaaggagaaggcatcttcaaaattgaggactcagcccaggtggccgg 1282
          |  |||     ||       ||||| ||||||    |   | | ||||     ||||| | |
Db    328 AAATGGGAAGACCGATCTGAGGGCGTCTTCAGGTTCTTGAAATCAGAGGCAGTGGCTCAG 269

Qy   1283 ctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctgagccgctccatc 1342
          || ||||| |     ||||||     | ||||| ||| || ||||| |||||  | ||
Db    268 CTATGGGGTAAAAAGAAGAACAACAGCAGCATGACCTATGAAAAGCTCAGCCGAGCTATG 209

Qy   1343 cgccagtattacaagaagggcat 1365
            |  | || ||||| |  |  ||
Db    208 AGATATTACTACAAAAGAGAAAT 186


RESULT  14
V15529
ID   V15529 standard; cDNA; 2975 BP.
AC   V15529;
DT   22-MAY-1998  (first entry)
DE   Human lymphoid-specific transcription factor NERF-1 cDNA.
KW   ets-related; human; lymphoid-specific transcription factor; NERF-1;
KW   screening assay; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             122. .1687
FT                   /*tag= a
FT                   /product= NERF-1
PN   US5721113-A.
PD   24-FEB-1998.
PF   03-JAN-1995; 368281.
PR   03-JAN-1995; US-368281.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Kunsch CA, Libermann TA, Oettgen JP;
DR   WPI; 98-168405/15.
DR   P-PSDB; W47237.
PT   NERF-1 and NERF-2 nucleic acids - encode ets-related human
PT   lymphoid-specific transcription factors
PS   Claim 1; Columns 17-22; 17pp; English.
CC   The present sequence encodes the ets-related lymphoid-specific
CC   transcription factor NERF-1, which can be used in screening assays
CC   for drugs that modulate NERF activity, and to treat patients having
CC   need of NERF-1. ets-related factors are involved in, e.g. cancer
CC   development, retrovirus replication, T-cell cytokine production,
CC   cell cycle effects, growth regulation and cell differentiation.
SQ   Sequence   2975 BP;     958 A;     575 C;     614 G;     828 T;


  Query Match             3.2%;  Score 60.4;  DB 1;  Length 2975;
  Best Local Similarity   53.9%;  Pred. No. 0.00056;
  Matches  124;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy   1151 gggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctat 1210
          ||  |  | | | |  ||||| |||| ||    ||  | || |  |      | |  | |
Db    518 GGAAACACAACCTATTTGTGGGAGTTTCTTTTAGATCTACTTCAAGATAAAAATACTTGT 577

Qy   1211 ggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcc 1270
            | | |  |||| |||    | |  || || ||||| |||||  | | ||| |||
Db    578 CCCAGGTATATTAAATGGACTCAGAGAGAAAAAGGCATATTCAAGCTGGTGGATTCAAAG 637

Qy   1271 caggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctg 1330
            ||  |   ||| ||||| |  |  ||||||   || | |||||||| || |   ||
Db    638 GCTGTCTCTAAGCTTTGGGGAAAGCATAAGAACAAACCAGACATGAACTATGAAACCATG 697

Qy   1331 agccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380
            | ||  |  |  |  | || ||| | | ||| ||  |    |||   ||
Db    698 GGACGAGCTTTGAGATACTACTACCAAAGGGGAATTCTTGCAAAGGTTGA 747
```

```
RESULT  15
V15530
ID   V15530 standard; cDNA; 3240 BP.
AC   V15530;
DT   22-MAY-1998  (first entry)
DE   Human lymphoid-specific transcription factor NERF-2 cDNA.
KW   ets-related; human; lymphoid-specific transcription factor; NERF-2;
KW   screening assay; ds.
OS   Homo sapiens.
FH   Key             Location/Qualifiers
FT   CDS             207. .1952
FT                   /*tag= a
FT                   /product= NERF-2
PN   US5721113-A.
PD   24-FEB-1998.
PF   03-JAN-1995; 368281.
PR   03-JAN-1995; US-368281.
PA   (HUMA-) HUMAN GENOME SCI INC.
PI   Kunsch CA, Libermann TA, Oettgen JP;
DR   WPI; 98-168405/15.
DR   P-PSDB; W47238.
PT   NERF-1 and NERF-2 nucleic acids - encode ets-related human
PT   lymphoid-specific transcription factors
PS   Claim 2; Columns 23-28; 17pp; English.
CC   The present sequence encodes the ets-related lymphoid-specific
CC   transcription factor NERF-2, which can be used in screening assays
CC   for drugs that modulate NERF activity, and to treat patients having
CC   need of NERF-2. ets-related factors are involved in, e.g. cancer
CC   development, retrovirus replication, T-cell cytokine production,
CC   cell cycle effects, growth regulation and cell differentiation.
SQ   Sequence   3240 BP;     1070 A;     592 C;     662 G;     916 T;


  Query Match             3.2%;  Score 60.4;  DB 1;  Length 3240;
  Best Local Similarity   53.9%;  Pred. No. 0.00057;
  Matches  124;  Conservative    0;  Mismatches  106;  Indels    0;  Gaps    0;

Qy   1151 gggcagcccatccacctgtggcagttcctcaaggagttgctactcaagccccacagctat 1210
          ||  |  | | | |  ||||| |||| ||    ||  | || |  |      | |  | |
Db    783 GGAAACACAACCTATTTGTGGGAGTTTCTTTTAGATCTACTTCAAGATAAAAATACTTGT 842

Qy   1211 ggccgcttcattaggtggctcaacaaggagaagggcatcttcaaaattgaggactcagcc 1270
            | | |  |||| |||    | |  || || ||||| |||||  | | ||| |||
Db    843 CCCAGGTATATTAAATGGACTCAGAGAGAAAAAGGCATATTCAAGCTGGTGGATTCAAAG 902

Qy   1271 caggtggcccggctgtggggcatccgcaagaaccgtcccgccatgaactacgacaagctg 1330
            ||  |   ||| ||||| |  |  ||||||   || | |||||||| || |   ||
Db    903 GCTGTCTCTAAGCTTTGGGGAAAGCATAAGAACAAACCAGACATGAACTATGAAACCATG 962

Qy   1331 agccgctccatccgccagtattacaagaagggcatcatccggaagccaga 1380
            | ||  |  |  |  | || ||| | | ||| ||  |    |||   ||
Db    963 GGACGAGCTTTGAGATACTACTACCAAAGGGGAATTCTTGCAAAGGTTGA 1012


Search completed: November 20, 1999, 21:06:19
Job time: 1640 sec
```